

# Application of reinforcement learning in mechatronic systems

First-Name Last-Name1, First-Name Last-Name2, First-Name Last-Name3

1 Affiliation, E-mail:

2 Affiliation, E-mail:

3 Affiliation, E-mail:

\*Corresponding E-mail:

**Abstract:** 本文目的是將相同的訓練算法能在 2D 環境上進行訓練和使用，將相同演算法套用到 3D 模擬環境進行訓練與使用，以驗證算法若將既有機電系統簡化成 2D 所測試的算法套用到 3D 模擬環境的可行性。將實體冰球機的機電系統簡化後導入 CoppeliaSim 模擬環境透過 Remote API 控制環境中的冰球機移動，OpenCV 來處理影像提供強化學習訓練和訓練後實際控制的輸入，強化學習訓練利用 OpenAI Gym 的 Pong Game 測試適合的訓練參數，再將算法套用到 CoppeliaSim 的場景中進行訓練。

**Keywords:** 類神經網路、強化學習、CoppeliaSim、OpenAI Gym

## Introduction

本文目的是將相同的訓練算法能在 2D 環境上進行訓練和使用，將相同演算法套用到 3D 模擬環境進行訓練與使用，以驗證算法若將既有機電系統簡化成 2D 所測試的算法套用到 3D 模擬環境的可行性。將實體冰球機的機電系統簡化後導入 CoppeliaSim 模擬環境透過 Remote API 控制環境中的冰球機移動，OpenCV 來處理影像提供強化學習訓練和訓練後實際控制的輸入，強化學習訓練利用 OpenAI Gym 的 Pong Game 測試適合的訓練參數，再將算法套用到 CoppeliaSim 的場景中進行訓練。

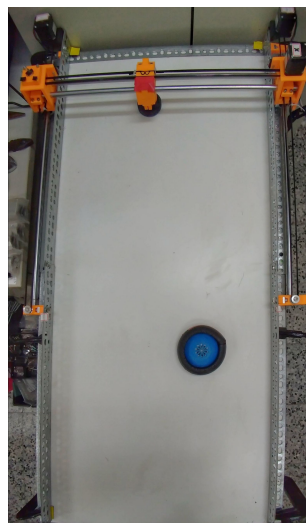


Figure 1: 實體的冰球機

## Literature Review

Atari Video Computer System 是 Atari Inc. 在 1977 年 9 月 11 日發行的一款家用遊戲機，直到 1982 年更名為 Atari 2600，Pong 是第一個主要發行的遊戲。OpenAI 是一個人工智慧研究實驗室，研究著重於強化學習，Gym 具有多種不同的環境，並以標準化定義環境，其中包括 Atari 2600 的遊戲、Classic control、Box2D 等環境供機器學習訓練使用。

冰球機與 Pong 都是透過玩家的眼睛觀察球的移動並透過操控擊錘進行攻擊與防守，並依照對手的球路來採取不同的策略。強化學習的訓練是透過影像處理，將擊錘及球的位置過濾出來，比較兩幀間擊錘與球的移動，透過強化學習判斷擊錘的移動，進行訓練。

## 設計流程

設計流程分為簡化模、型機器學習訓練、模擬環境訓練。

簡化模型將冰球機簡化成為一個自由度，簡化後與 Pong 的操控相同，並分別透過 2D 遊戲 Pong-v0 與 3D 模擬環境 CoppilaSim 供強化學習訓練。

機器學習訓練透過影像處理，過濾出擊錘與球，並將兩幀影像進行比較，透過調整環境參數與決策參數的權重，環境參數使用 ReLU Function 進行優化，決策參數則是透過 Log probability 和 Softmax 進行運算並在每局結束後透過 back-propagation 修正與更新偏差，藉由不斷的學習持續進步。

模擬環境訓練在 CoppilaSim 中，抓取場景中的攝影機，